



اَبُو سَيِّدِي نَبِي كَوْنِي مَارَا  
UNIVERSITI  
TEKNOLOGI  
MARA

# THE DOCTORAL RESEARCH ABSTRACTS

Volume: 7, Issue 7 May 2015

## SEVENTH ISSUE

INSTITUTE of GRADUATE STUDIES

*Leading You To Greater Heights, Degree by Degree*

IGS Biannual Publication

Name :

**Nuru'l - 'Izzah Othman**

Title :

**Computational Model of Stepwise Quantitative Assessment of Equations Correctness**

Supervisor :

**Prof. Dr. Arsmah Ibrahim (MS)**

**Prof. Dr. Zainab Abu Bakar (CS)**

Existing computational solutions for stepwise correctness checking of free-response solution schemes consisting of equations only consider providing qualitative feedbacks. Hence, this research intends to propose a computational model of a typical stepwise correctness checking of a scheme of student-constructed responses normally (usually) performed by a human examiner with the provision of quantitative feedbacks. The responses are worked solutions on solving linear algebraic equations in one variable. The proposed computational model comprises of computational techniques of key marking processes, and has enabled a marking engine prototype, which has been developed based on the model, to perform stepwise correctness checking and scoring of the response of each step in a working scheme and of the working scheme as well. The assigned numeric score of each step, or analytical score, serves as a quantitative feedback to inform students on the degree of correctness of the response of a particular step. The numeric score of the working scheme, or overall score, indicates the degree of correctness of the whole working scheme. Existing computational solutions that are currently available determine response correctness based on mathematical equivalence of expressions. In this research, the degree of correctness of an equation is based on the structural identicalness of the constituting mathtokens, which is evaluated using a correctness measure formulated in this research. The experimental verification shows that the evaluation of correctness by the correctness measure is comparable to human judgment on correctness. The computational model is formalized mathematically by basic concepts from Multiset Theory, while the process framework is supported by basic techniques and processes from the field of textual information retrieval that have been adapted to suit the problem of this research. The data used are existing worked solutions on solving linear algebraic equations in one variable from a previous pilot study as well as new sets of responses that were collected using the marking engine prototype. The experimental test of correctness shows that the computational model is able to generate the expected output. Hence, the underlying computational techniques of the model can be regarded as correct. The agreement between the automated and the manual marking methods were analysed in terms of the agreement between the correctness scores. The method agreement analyses were conducted in two phases. The analysis in Phase I involved a total of 561 working schemes which comprised of 2021 responses and in Phase II a total of 350 working schemes comprising of 1385 responses were used. The analyses involved determining the percent agreement, degree of correlation and degree of agreement between the automated and manual scores. The accuracy of the scores was determined by calculating the average absolute errors present in the automated scores, which are calibrated by the average mixed errors. The results show that both the automated analytical scores and the automated overall scores exhibited high percent agreement, high correlation, high degree of agreement and small average absolute and mixed errors. It can be inferred that the automated scores are comparable with manual scores and that the stepwise correctness checking and scoring technique of this research agrees with the human marking technique. Therefore, the computational model of stepwise quantitative assessment is a valid and reliable model to be used in place of a human examiner to check and score responses to similar questions used in this research for both formative and summative assessment settings..